Title: **Identifying molecular features that distinguish fluvastatin-sensitive breast tumor cells**

Authors: Carolyn A. Goard[1,2], Michelle Chan-Seng-Yue[1,3], Peter J. Mullen[1], Ariel D. Quiroga[4], Amanda R. Wasylishen[1,2], James W. Clendening[1,2], Dorota H. S. Sendorek[3], Syed Haider[1], Richard Lehner[4], Paul C. Boutros[2,3], Linda Z. Penn[1,2]

Affiliations: [1]Ontario Cancer Institute and Campbell Family Institute for Cancer Research, Princess Margaret Cancer Centre, University Health Network, and [2]Department of Medical Biophysics, University of Toronto, 610 University Avenue, Toronto, Ontario, M5G 2M9, Canada; [3]Informatics and Biocomputing Platform, Ontario Institute of Cancer Research, 101 College Street, Toronto, Ontario, M5G 0A3, Canada; [4]Department of Pediatrics and Cell Biology, Group on Molecular and Cell Biology of Lipids, University of Alberta, 328 Heritage Medical Research Centre, Edmonton, Alberta, T6G 2S2, Canada.

Short Title: Predictors of fluvastatin sensitivity in breast cancer

Corresponding Author: **Dr. Linda Z. Penn**

Ontario Cancer Institute
610 University Avenue, Room 9-625
Toronto, Ontario, M5G 2M9, Canada.

Phone: (416) 946-2276
Fax: (416) 946-2840
E-mail: lpenn@uhnres.utoronto.ca

SUPPLEMENTARY METHODS:

Cell proliferation assay:

10,000 cells/well of a 12-well plate were seeded in duplicate for each of five days. Cells were trypsinized and counted by hemocytometer. Doubling times were calculated from an exponential growth curve using GraphPad Prism (v5.0).

Measurement of cholesterol and lipids:

1,000,000 cells were seeded overnight on 10 cm plates and treated with ethanol vehicle control, 10 μM fluvastatin or serum withdrawal to 0.01% FBS for 16 h. Cells were washed with PBS, harvested on ice in 1.5 ml ice-cold PBS supplemented with "cOmplete" protease inhibitor cocktail (Roche) and processed similarly to our previous report [1]. Cells were sonicated, phospholipid polar heads were removed by incubation with Phospholipase C, and total lipids were extracted in Folch solution (chloroform:methanol, 2:1, [v/v]). Lipids were derivatized and analyzed on an Agilent 6890 gas chromatographer with a flame-ionization detector and a Zebron ZB-5 column (Phenomenex; 30 m x 0.32 mm I.D. x 0.25 μm film thickness). Results were analyzed with Agilent Chemstation software and corrected to a 2 μg tridecanoin internal standard added during lipid extraction.

Data mining for fluvastatin sensitivity-associated gene expression:

Publicly available microarray data for 51 breast cancer cell lines [2] were preprocessed using updated cdf files that map probes to Entrez Gene IDs (hgu133ahsentrezgcdf; v15.0.0) and the RMA algorithm [3] as implemented in the affy package (v1.32.1) for R (v2.14.2). Of the cell lines used in this report, HCC1806, SW527 and HCC1419 were not represented in this dataset. A Y chromosome filtration was applied to remove unexpressed genes according to background hybridization levels [4]. Genes with intensities below 5.0 in all cell lines were excluded from downstream analyses. An analysis of Gene Ontology (GO) term enrichment was performed on the dataset of fluvastatin sensitivity-associated genes using the web-based, high-throughput version of GOMiner (Build 291) [5]. All GO terms were evaluated in human-specific

databases, and statistical constraints for summary reports were set to 0.1. One thousand randomizations were utilized.

Verification of mRNA abundance correlations with NanoString technology:

Forty-eight samples (23 cell lines in duplicate, 2 cell lines as single samples) were processed on the NanoString platform to detect the mRNA abundance of 25 genes, according to standard protocols at the University Health Network Microarray Centre (Toronto, Canada). NanoString data was preprocessed using the NanoStringNorm (v1.1.11) package for R (v2.15.2) [6]. The geometric mean was used to normalize for technical assay variation while the background count level was estimated using the maximum of the negative controls. Finally, RNA content was normalized using the geometric mean of three housekeeping genes, glucuronidase, beta (*GUSB*), glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) and hypoxanthine phosphoribosyltransferase 1 (*HPRT1*). One replicate of HS578T was removed from the data processing due to positive normalization outside of the acceptable range. For sample replicates, the mean for each gene was calculated and utilized for downstream analysis. Affymetrix data was processed as noted above. Affymetrix expression levels for the 25 genes from the samples evaluated on NanoString were isolated to create a comparable dataset. Spearman correlation coefficients between the NanoString and Affymetrix expression data were calculated. Both datasets were also independently correlated to the mean fluvastatin MTT_{50} of each cell line, calculating Pearson correlation coefficients. Samples were then compared using the diana clustering algorithm found in the cluster package (v1.14.3) for R.(v2.15.2). All visualizations were created using the lattice (v0.20-10) and latticeExtra (v1.26.0) packages in R (v2.15.2).

SUPPLEMENTARY REFERENCES:

1. Clendening JW, Pandyra A, Li Z, Boutros PC, Martirosyan A, Lehner R, Jurisica I, Trudel S, Penn LZ (2010) Exploiting the mevalonate pathway to distinguish statin-sensitive multiple myeloma. Blood 115:4787-4797. doi: 10.1182/blood-2009-07-230508

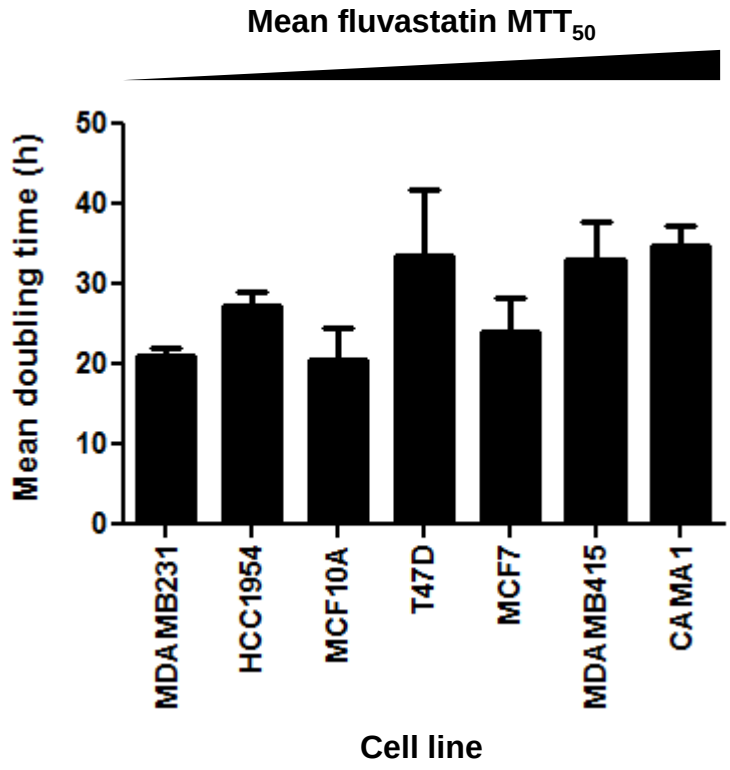
2. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, Clark L, Bayani N, Coppe JP, Tong F, Speed T, Spellman PT, DeVries S, Lapuk A, Wang NJ, Kuo WL, Stilwell JL, Pinkel D, Albertson DG, Waldman FM, McCormick F, Dickson RB, Johnson MD, Lippman M, Ethier S, Gazdar A, Gray JW (2006) A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell 10:515-527.

3. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4:249-264. doi: 10.1093/biostatistics/4.2.249

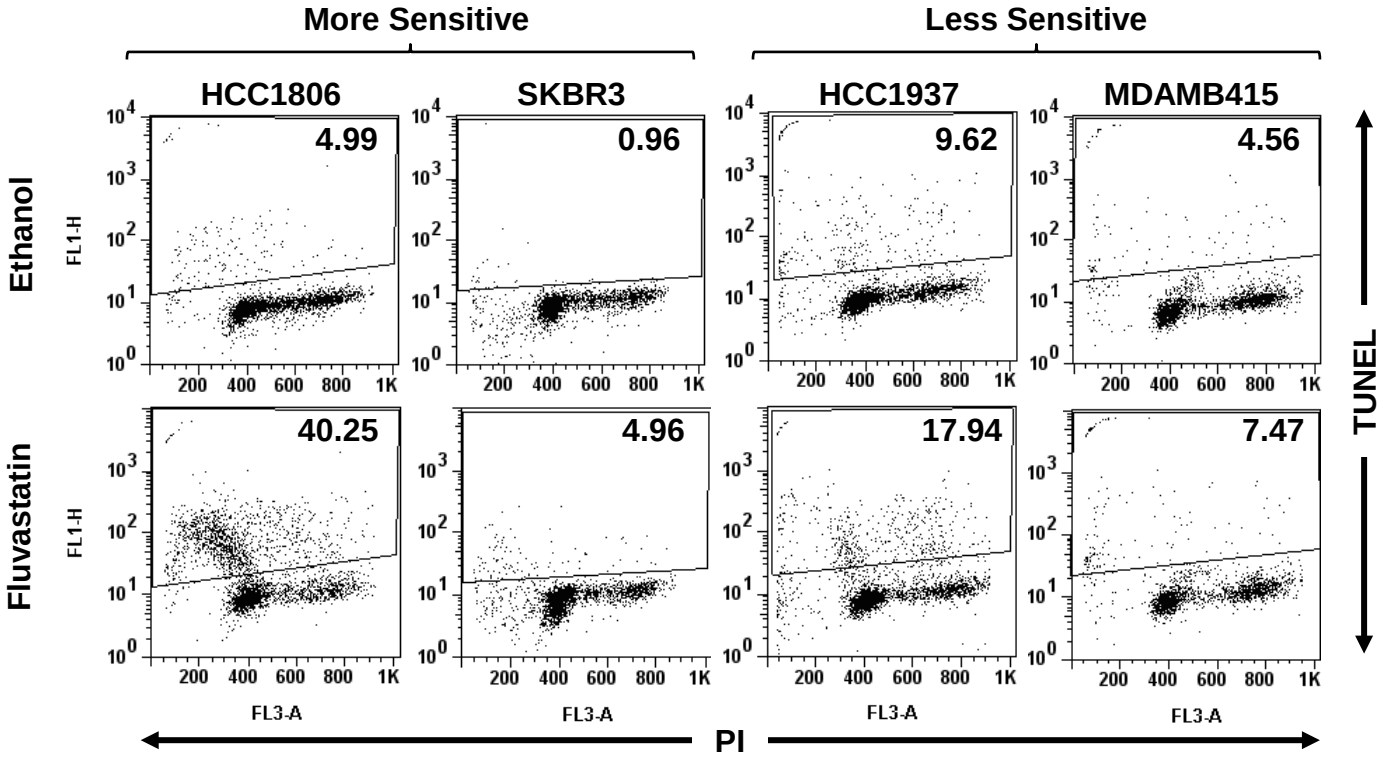
4. Bachtiary B, Boutros PC, Pintilie M, Shi W, Bastianutto C, Li JH, Schwock J, Zhang W, Penn LZ, Jurisica I, Fyles A, Liu FF (2006) Gene expression profiling in cervical cancer: an exploration of intratumor heterogeneity. Clin Cancer Res 12:5632-5640. doi: 10.1158/1078-0432.CCR-06-0357

5. Zeeberg BR, Qin H, Narasimhan S, Sunshine M, Cao H, Kane DW, Reimers M, Stephens RM, Bryant D, Burt SK, Elnekave E, Hari DM, Wynn TA, Cunningham-Rundles C, Stewart DM, Nelson D, Weinstein JN (2005) High-Throughput GoMiner, an 'industrial-strength' integrative gene ontology tool for interpretation of multiple-microarray experiments, with application to studies of Common Variable Immune Deficiency (CVID). BMC bioinformatics 6:168. doi: 10.1186/1471-2105-6-168

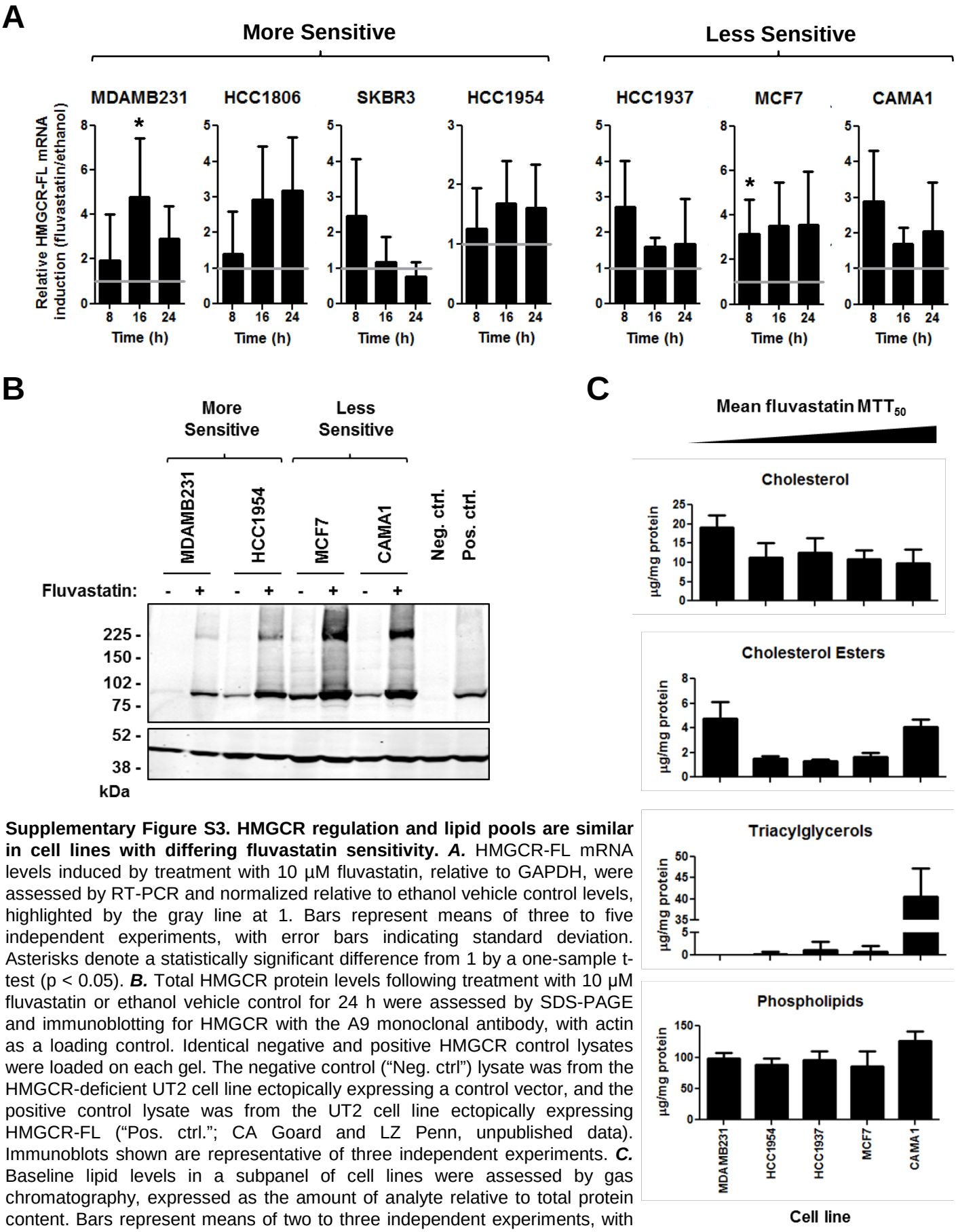
6. Waggott D, Chu K, Yin S, Wouters BG, Liu FF, Boutros PC (2012) NanoStringNorm: an extensible R package for the pre-processing of NanoString mRNA and miRNA data. Bioinformatics 28:1546-1548. doi: 10.1093/bioinformatics/bts188



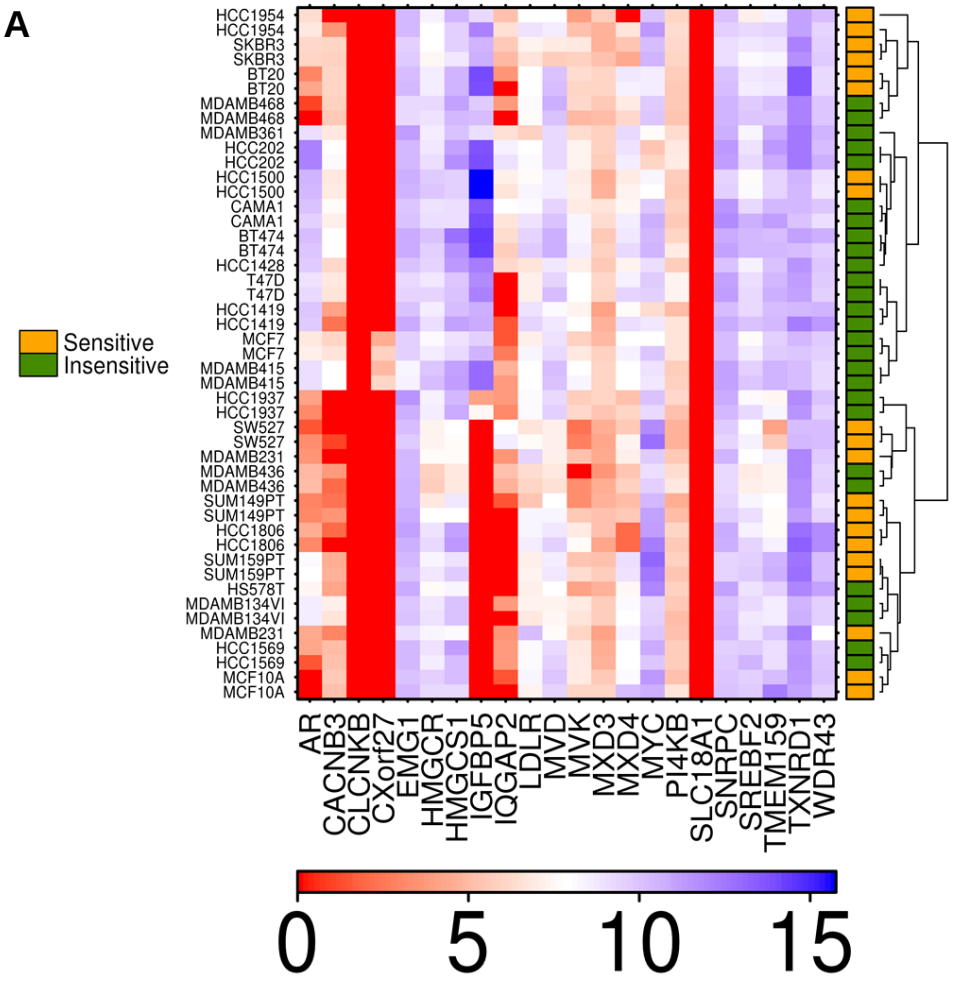
Supplementary Figure S1. Fluvastatin sensitivity does not appear to be solely associated with proliferative rate in a subpanel of breast cell lines. Subconfluent cells were seeded in duplicate and counted in duplicate over five days by hemocytometer. Doubling times were calculated from exponential growth curves. Bars represent mean doubling times from three independent experiments, with error bars indicating standard deviation.

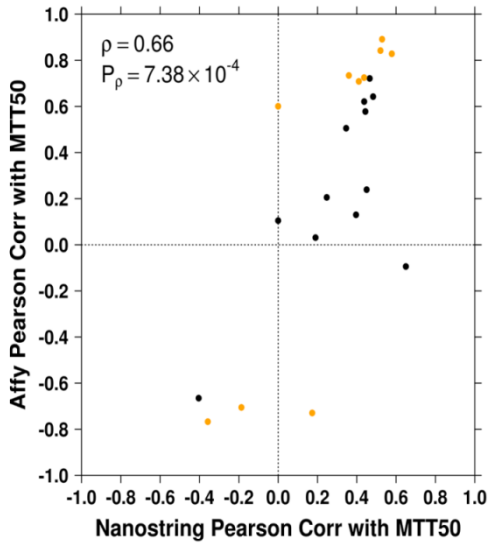


Supplementary Figure S2. Representative TUNEL vs. PI dotplots from additional cell lines demonstrating differential induction of apoptosis following fluvastatin treatment. The proportion of cells undergoing apoptosis following treatment with fluvastatin or ethanol vehicle control for 72 h was determined by TUNEL staining and flow cytometric detection. Numbers highlighted in the upper right corners of each dotplot represent the number of events detected in the TUNEL-positive gate shown. Experiments were repeated three times.



Supplementary Figure S3. HMGCR regulation and lipid pools are similar in cell lines with differing fluvastatin sensitivity. ***A.*** HMGCR-FL mRNA levels induced by treatment with 10 µM fluvastatin, relative to GAPDH, were assessed by RT-PCR and normalized relative to ethanol vehicle control levels, highlighted by the gray line at 1. Bars represent means of three to five independent experiments, with error bars indicating standard deviation. Asterisks denote a statistically significant difference from 1 by a one-sample t-test ($p < 0.05$). ***B.*** Total HMGCR protein levels following treatment with 10 µM fluvastatin or ethanol vehicle control for 24 h were assessed by SDS-PAGE and immunoblotting for HMGCR with the A9 monoclonal antibody, with actin as a loading control. Identical negative and positive HMGCR control lysates were loaded on each gel. The negative control ("Neg. ctrl") lysate was from the HMGCR-deficient UT2 cell line ectopically expressing a control vector, and the positive control lysate was from the UT2 cell line ectopically expressing HMGCR-FL ("Pos. ctrl."; CA Goard and LZ Penn, unpublished data). Immunoblots shown are representative of three independent experiments. ***C.*** Baseline lipid levels in a subpanel of cell lines were assessed by gas chromatography, expressed as the amount of analyte relative to total protein content. Bars represent means of two to three independent experiments, with error bars indicating standard deviation.





Supplementary Figure S4. mRNA abundance of a subset of genes is similar mined Affymetrix array data and NanoString detection. Baseline mRNA levels of 25 genes were assessed in a panel of 25 breast cell lines and compared to mRNA abundance in these cell lines as determined by mining publicly available Affymetrix array-based data. ***A.*** Hierarchical Diana clustering confirms that duplicate samples of cell lines are most similar to each other in mRNA abundance patterns, with the exception of one cell line (MDAMB231). Heatmap displays normalized mRNA abundance values. ***B.*** A comparison of the correlations between mRNA abundance of each gene determined by either Affymetrix array data or NanoString verification and fluvastatin MTT_{50} demonstrates that similar correlations are detected for most genes. Yellow data points indicate the 10 genes comprising the gene signature while the black data points represent the remaining genes processed on NanoString. The Spearman correlation coefficient, ρ, and its associated *p*-value demonstrate a strong concordance between associations of mRNA abundance detected by either method with fluvastatin sensitivity.